A Large Ensemble of F-theory Geometries:

the Weak, the Strong, and the Non-Higgsable

Jim Halverson
Northeastern University

Based on:

1706.02299 with **Cody Long** and Ben Sung
1707.00655 with Jon Carifio, Dima Krioukov, and Brent Nelson
1710.09374 with **Cody Long** and Ben Sung
1711.06685 with Jon Carifio, Will Cunningham, Dima Krioukov, **Cody Long**, and Brent Nelson

A Large Ensemble of F-theory Geometries:

the Weak and the Strong via AI Agents, and the Non-Higgsable

Jim Halverson
Northeastern University

Based on:

1706.02299 with **Cody Long** and Ben Sung
1707.00655 with Jon Carifio, Dima Krioukov, and Brent Nelson
1710.09374 with **Cody Long** and Ben Sung
1711.06685 with Jon Carifio, Will Cunningham, Dima Krioukov, **Cody Long**, and Brent Nelson

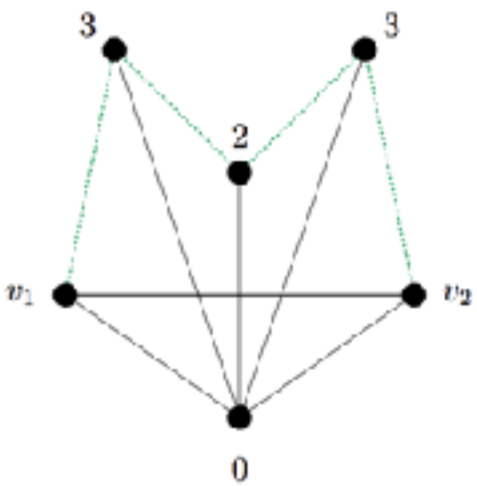
To appear: work with Nelson, **Ruehle**, and work with Long, **Ruehle**, and Tian

Outline:

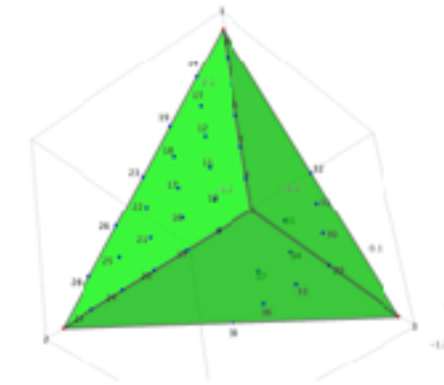
What does the F-theory landscape look like?

- A **Large** Ensemble of F-theory Geometries
- **Universality** of NHC and Strong Coupling
- **Anomaly** Detection: an E6 warmup and exploring the IIB lamppost with reinforcement learning (AI).

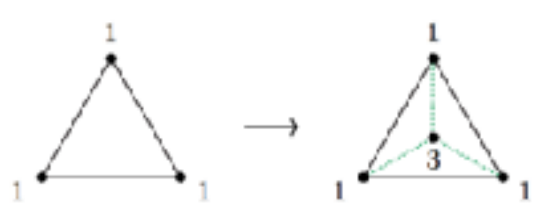
More Universality Evidence: Taylor-Wang

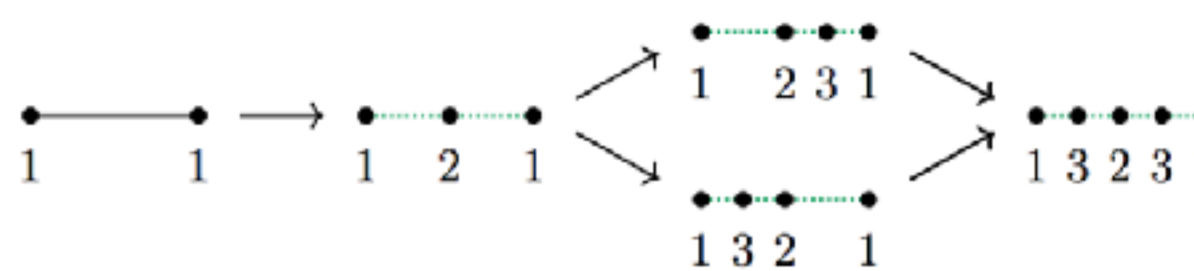


A Large Ensemble of F-theory Geometries



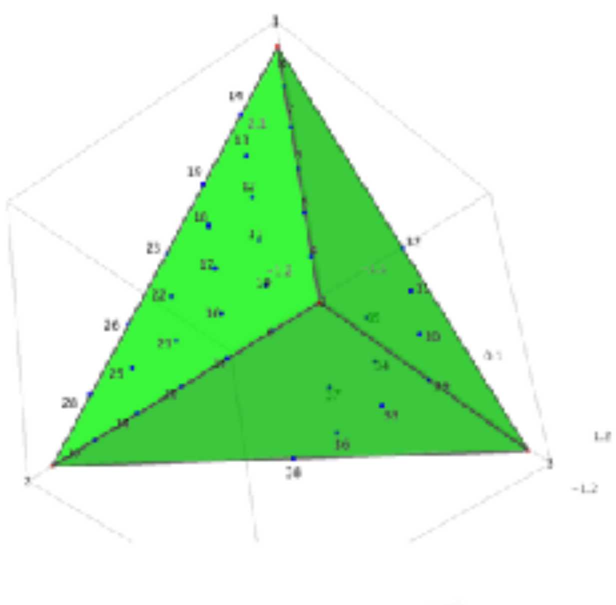
296231521306658847673529878342496866080378021782076722241462425
150065298508845621155695644076004524635389132997146097067111728
163388031396056782083347614866566976655879021984012793769157213
779169816753325335978325856186523598065404800375507248464342388
511047235006655822119955724586948796896518847577026036417026840
946433790956016724443830558425085959438371548897609540032670520
282917998350399793115958231423494399655849101565013413100994870
230661418306662241733727360211923208425845972788270516755913617
371792570199269731190754783864008989335957900232620795965279395
190240666076073359738714472109475716934497800719015826299673405
211214139176513510727323709162299991669746741171716096000000000
000000000000000000000000000000000000000000000000000000000000000

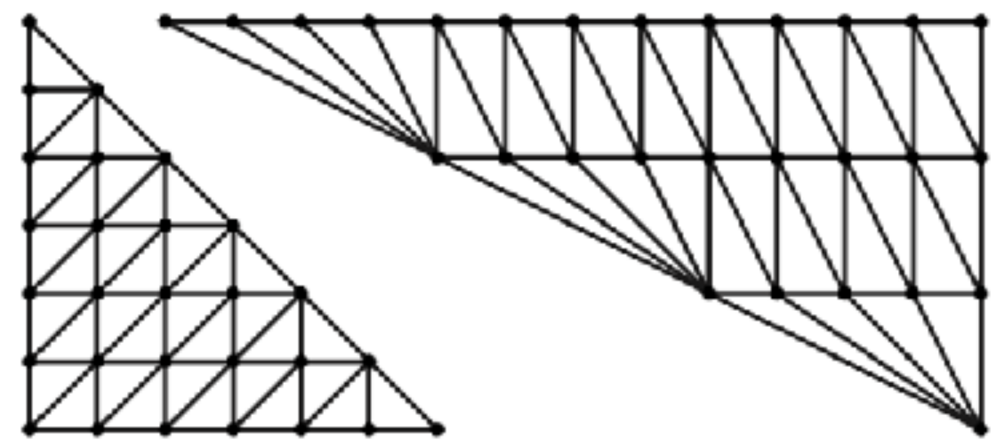




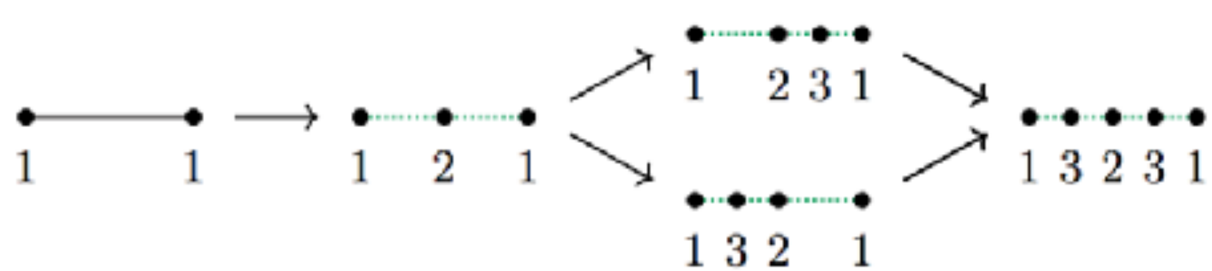
A Large Ensemble: How?

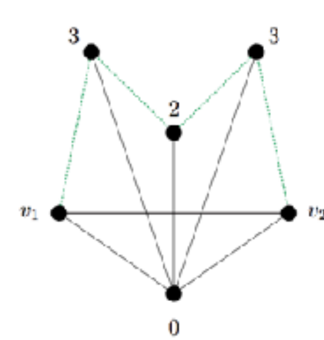
Starting point: weak Fano toric threefold ***base***



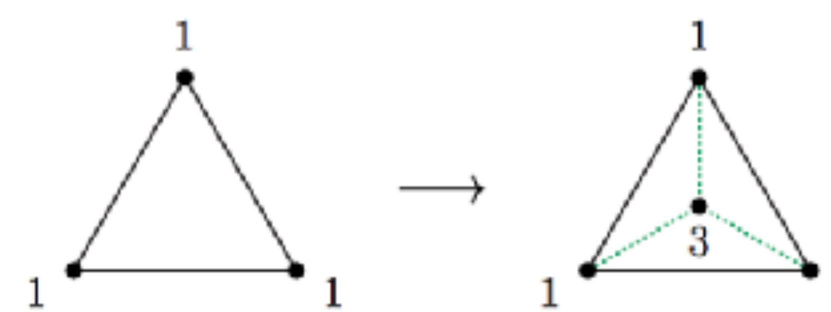


Build on top: sequences of toric blowups





Curve / “Edge” Blowups



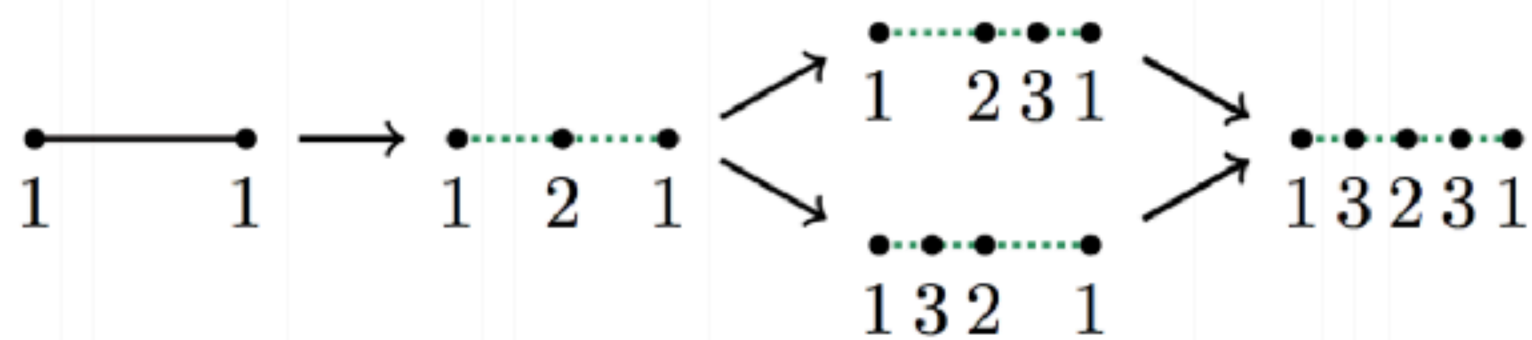
Point / “Face” Blowups

Language for the Blowups

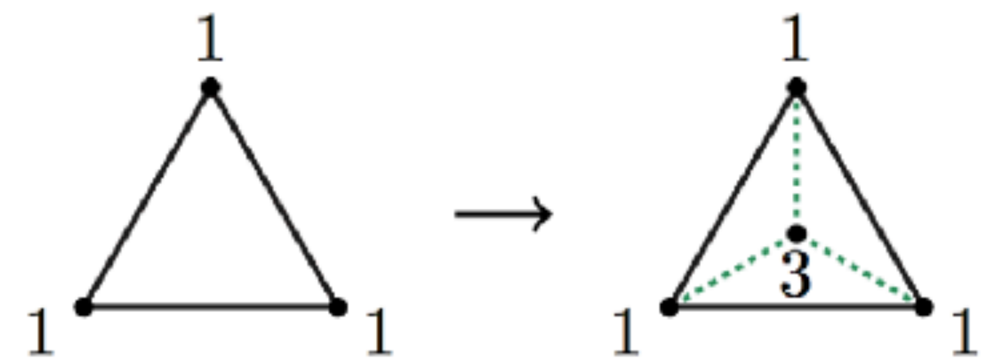
- We’ll refer to a sequence of blowups as a **“tree”**
- exceptional divisor from the sequence is a **“leaf”**
- Trees over edges = **“edge trees”**
- Trees over faces = **“face trees”**
- Points on polytope = leaves on ground = **“roots”**
- **Classify all trees** with h ≤ 6 for all leaves.
- Do so by exhaustively constructing the toric blowups.

Classification of Trees

- All 5 $h \leq 3$ edge trees.



- Both $h \leq 3$ face trees.

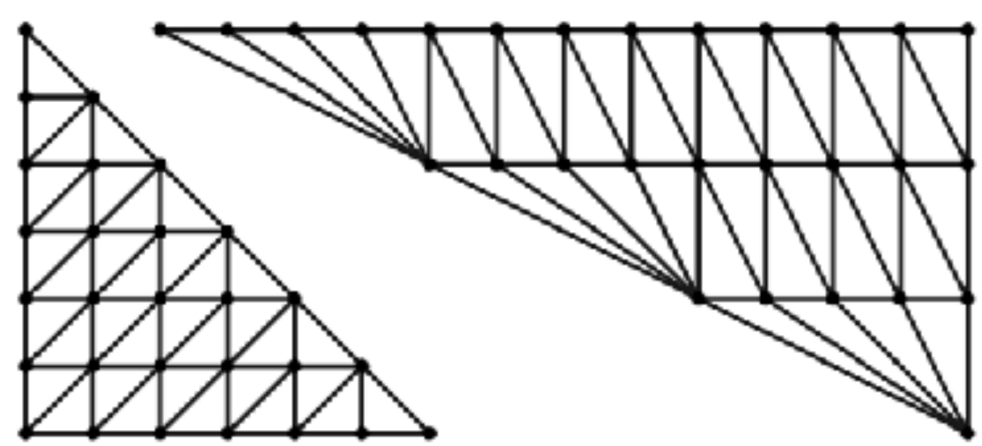


- # for $h \leq N$:

N	# Edge Trees	# Face Trees
3	5	2
4	10	17
5	50	4231
6	82	41,873,645

Forests from Trees

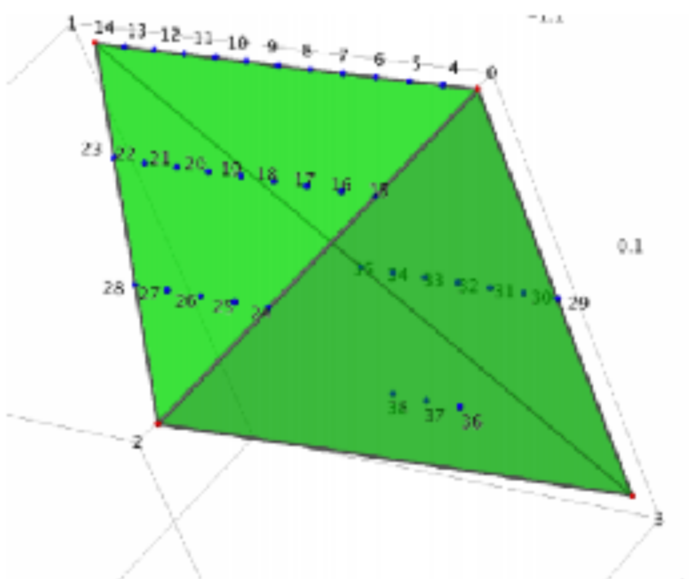
- Each "tree" is data representing a local sequence of blowups.
- **Form "forest"** (threefold base B) from trees by systematically adding trees to FRST of a 3d reflexive polytope. Face trees first, then edge.
- **Count:** polytopes whose FRST's have the largest number of faces and edges dominate the ensemble.
- **Two polytopes dominate:** have 108 edges and 72 faces, very large facet.

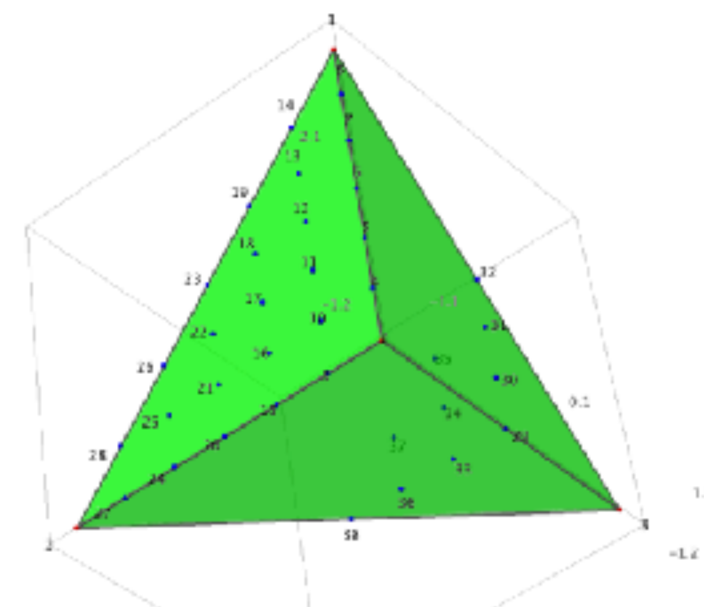


large face of Δ_1° **large face of** Δ_2°

The Large Ensemble

The dominant polytopes:





Each has 108 toric curves (edges) and 72 toric points (triangles) when triangulated. The number of bases in these ensembles is:

$$|S_{\Delta_1^\circ}| = \frac{2.96}{3} \times 10^{755} \qquad |S_{\Delta_2^\circ}| = 2.96 \times 10^{755}$$

Studied network properties: [Carifio, Cunningham, JH, Krioukov, Long, Nelson]

Universality of NHC and Strong Coupling

(technique: universality from precise construction algorithm)

Results:

- likelihood of NHC (including type II)
- high probability minimal geometric gauge groups
- the Sen limit almost never exists (F-theory is F-theoretic!)

Non-Higgsable 7-branes

Some selective progress: Anderson, Braun, del Zotto, Halverson, Heckman, Grassi, Morrison, Schafer-Nameki, Shaneson, Taylor, Vafa, Wang

F	a	b	c	Sing.	G
I_0	≥ 0	≥ 0	0	none	none
I_n	0	0	$n \geq 2$	A_{n-1}	$SU(n)$ or $Sp(\lfloor n/2 \rfloor)$
II	≥ 1	1	2	none	none
III	1	≥ 2	3	A_1	$SU(2)$
IV	≥ 2	2	4	A_2	$SU(3)$ or $SU(2)$
I_0^*	≥ 2	≥ 3	6	D_4	$SO(8)$ or $SO(7)$ or G_2
I_n^*	2	3	$n \geq 7$	D_{n-2}	$SO(2n-4)$ or $SO(2n-5)$
IV^*	≥ 3	4	8	E_6	E_6 or F_4
III^*	3	≥ 5	9	E_7	E_7
II^*	≥ 4	5	10	E_8	E_8

$$f = \tilde{f} \prod_i x_i^{a_i}$$

$$g = \tilde{g} \prod_i x_i^{b_i}$$

$$\Delta = \tilde{\Delta} \prod_i x_i^{c_i}$$

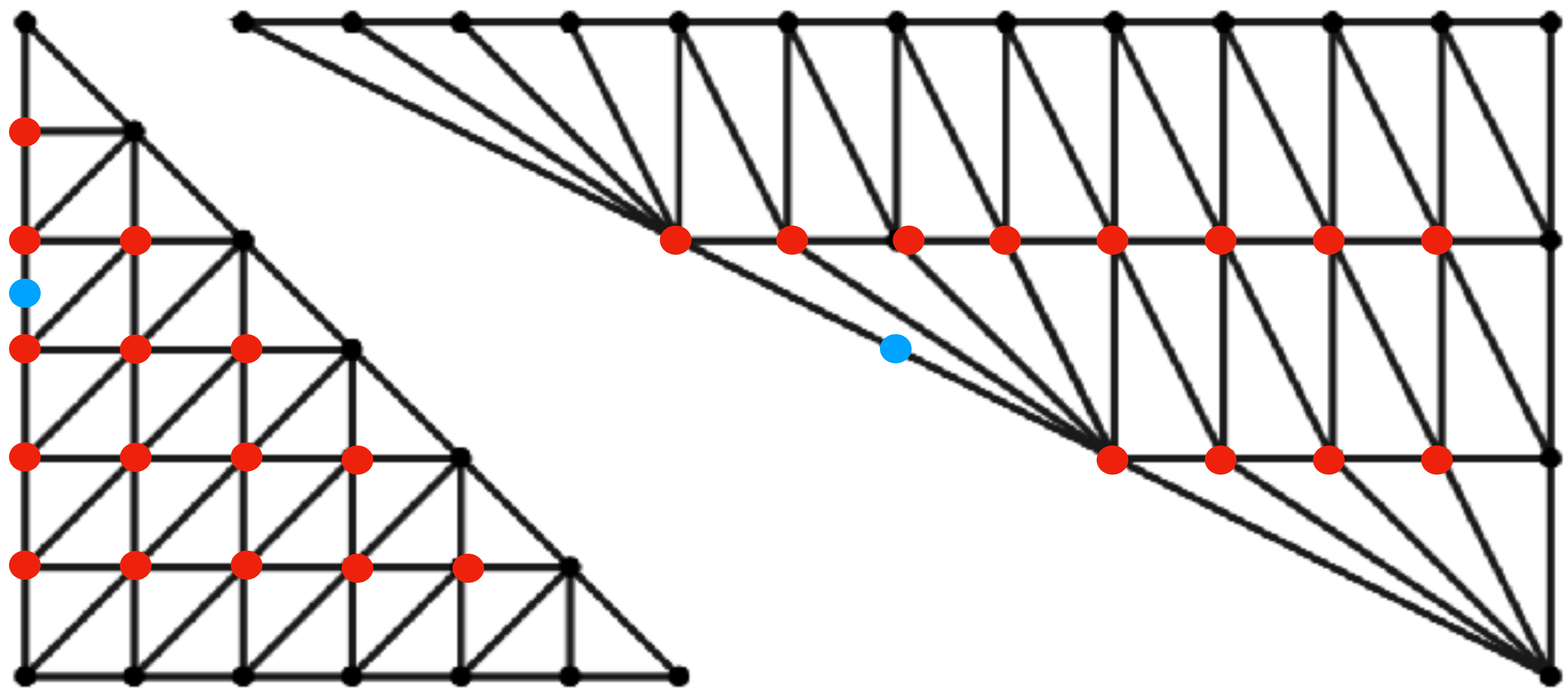
$$c_i = min(3a_i, 2b_i)$$

- **Non-Higgsable seven-brane:** $c_i > 0$ for some i. (NH7)
  $$G \in \{E_8, E_7, E_6, F_4, SO(8), SO(7), G_2, SU(3), SU(2)\}$$
- **Cannot be Higgsed** by a complex structure deformation.
- **Non-Higgsable clusters:** network of intersecting NH7. (NHC).
- **Entirely determined by topology of B!**

Universality of NH7 JH, Long, Sung

- Consider an edge or facet of a polytope, and perform a height > 2 blowup on that edge or facet.

- This cuts out a special monomial in **f**, **g**, forces type II NH7 on all divisors correspondir



blow up

type II NH7

All face trees (except for one on ground) have a h > 2 leaf.

All but two edge trees have a h > 2 leaf.

$$P(\mathrm{NHC\ in\ } S_{\Delta_1^\circ}) \geq 1 - 1.01 \times 10^{-755}$$

$$P(\mathrm{NHC\ in\ } S_{\Delta_2^\circ}) \geq 1 - .338 \times 10^{-755}$$

Universality of Large Gauge Sectors

JH, Long, Sung

- E_8 on roots (divisors on facet) are extremely common.

- **Theorem:** A leaf built on E_8 roots with height h = 1,2,3,4,5,6 has Kodaira fiber $F = II^*, IV^*_{ns}, I^*_{0ns}, IV_{ns}, II, -$ and geometric gauge group $E_8, G_2, SU(2), -, -$ respectively.

- Let H_i be number of height i leaves above E_8 roots. Then:

$$G \geq E_8^{10} \times F_4^{18} \times U^9 \times F_4^{H_2} \times G_2^{H_3} \times A_1^{H_4} \qquad U \in \{G_2, F_4, E_6\}$$
$$rk(G) \geq 160 + 4H_2 + 2H_3 + H_4$$

with probability $\geq .999995$

- Raises interesting cosmological question — dark glueballs?

Universality of Strong Coupling

(Sen limit almost never exists). **JH, Long, Sung**

- **Sen's limit:** weakly coupled IIb limit in CS moduli space.
- **Does not exist** if you have NH7 with fiber more singular than I_0^* (i.e. no seven-branes with exceptional G at weak coupling.)
- Such NH7 exist if either of the following are satisfied:
  - two h=2 and one h=3 curve blowups on same polytope edge.
  - h=3 point blowup strictly interior to polytpe face.
- Therefore **prob. of a Sen limit exists < 3 x 10^{-391}** in this ensemble.
- **Conclusion?** F-theory is F-theoretic.
  Weakly coupled IIB lamppost effect could be very severe.

  Can we better understand the IIB lampost and its boundary?

Anomaly Detection

Q: how do you detect and understand rare phenomena
in relation to a much larger ensemble?

(example: monitor network for malicious activity.
Normal activity vs. amateur hackers vs. professional hackers)

Warmup: an E6 Puzzle

- **Gauge group result:** dominated by $G_i \in \{E_8, F_4, G_2, A_1\}$

  (interesting: groups with only self-conjugate reps!)

- **Something SM-useful?** E6? SU(3)?

  - Simple conditions / probabilities for them not known. **JH, Long, Sung**
  - In random samples, prob ~ 1/1000. **this is the anomaly**
  - When E6 arises in RS, on a distinguished vertex: (1,-1,-1).

- **Machine Learning:** **Carifio, Halverson, Krioukov, Nelson**

  Q: Can we train a ML model to accurately predict yes or no for E6 on (1,-1,-1)?

  Q: If so, can we learn how it makes its decision?

  in our paper: called **conjecture generation**.
  as a CS buzzword: **intelligible AI**.

Point: by using machine learning to generate conjectures, we may be able to take its numerical / empirical results and turn it into rigorous results.

Training / Evaluating the Model

- **Supervised machine learning:** given a large number of (input,output) pairs, **learn to predict** output given input, and then **test on unseen data**, see how well the model does.

- **Training data:** 10000 random with no E6, 10000 random with E6.
input: related to number of leaves of certain “heights above.”
output: whether or not E6 on (1,-1,-1)

- **Evaluation:** models work well, > 99% accuracy. (see next slide)

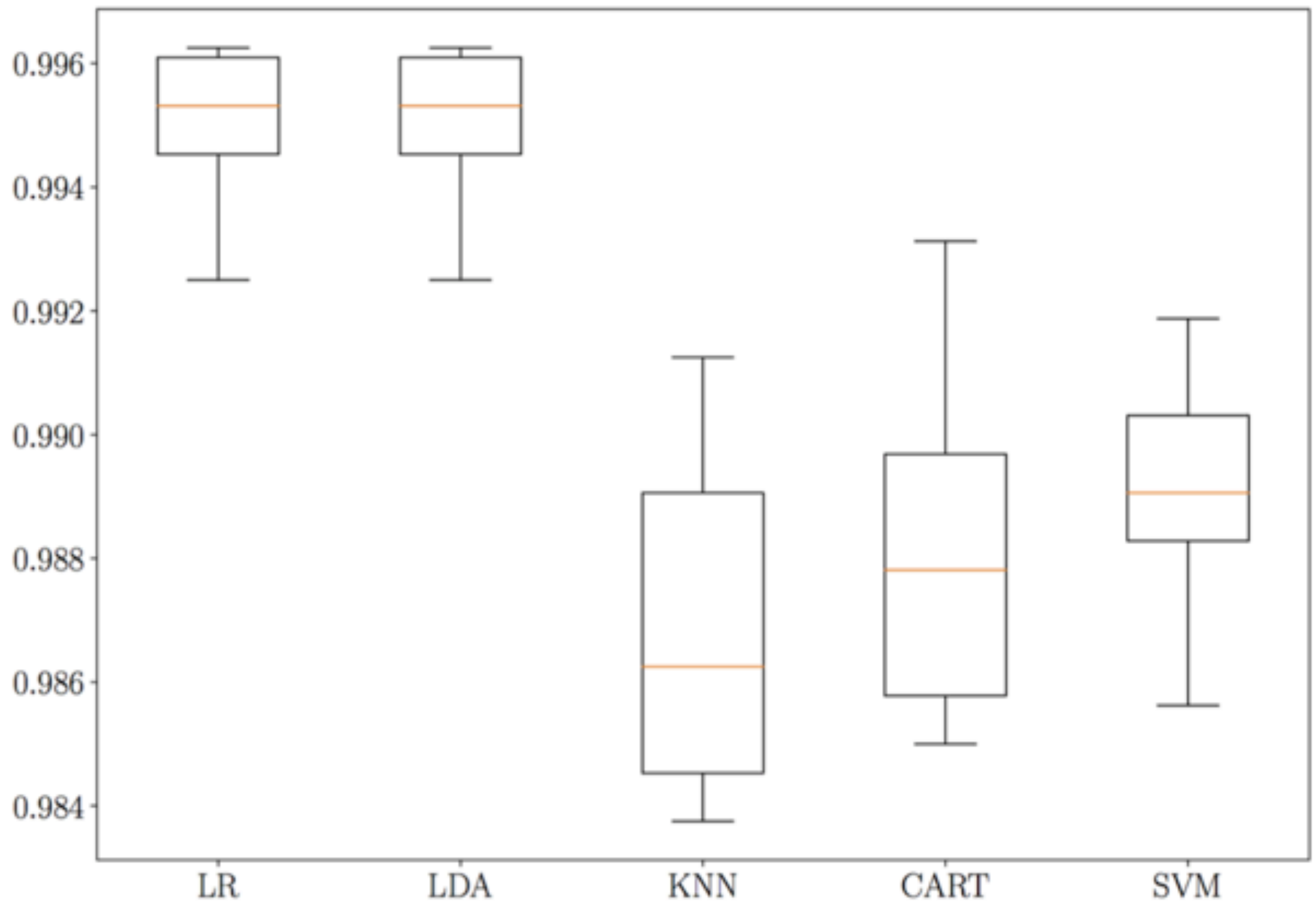
- **Conjecture generation:** model shows only two variables matter, leads to conjecture and theorem, use to compute prob(E6).

$$P(E_6 \text{ on } v_{E_6} \text{ in } T) = \left(1 - \frac{36}{82}\right)^9 \left(1 - \frac{18}{82}\right)^9 \simeq .00059128$$

From Theorem : $.00059128 \times 2 \times 10^6 = 1182.56$
From Random Samples : $1183, 1181, 1194, 1125, 1195$

Rigorous result from ML

Evaluating the Model on Unseen Data



	LR	LDA	KNN	CART	SVM
50/50 Validation Set	.994	.994	.982	.987	.989
Unenriched Set	.988	.988	.981	.988	.983.

- **Displayed:** whisker plots of % accuracy with 10-fold cross validation.
- **Gold bar:** mean % accuracy.
- **Factor analysis:** only two of the variables really matter:

$$(a_{max}, |S_{a_{max}, v_{E6}}|)$$

Anomaly Detection: Weak Coupling?

First theorem: due to  

pretty good, but didn't make paper.

Second theorem: due to Cody 

very good, this is in the paper.

Third (strongest?) theorem due to 

Goal: not just probability, but ***explore*** WC lamppost and probe its boundary.

JH, Long, Ruehle, Tian

Reinforcement Learning

supervised ML **predicts**, RL (AI) **explores / searches**

most famous examples: (?) AlphaGo & AlphaGo Zero

in strings: see Halverson, Nelson, Ruehle to appear soon.

- an ***agent*** interacts in an ***environment***.
- it perceives a ***state*** from ***state space***.
- its ***policy*** picks and executes an action, given the state.
- agent arrives in new state, receives a ***reward.***
- succesive rewards accumulate into ***return.***
- return may penalize future rewards via ***discount factor***.
- policy optimized to maximize reward, i.e. **agent learns how to act!**

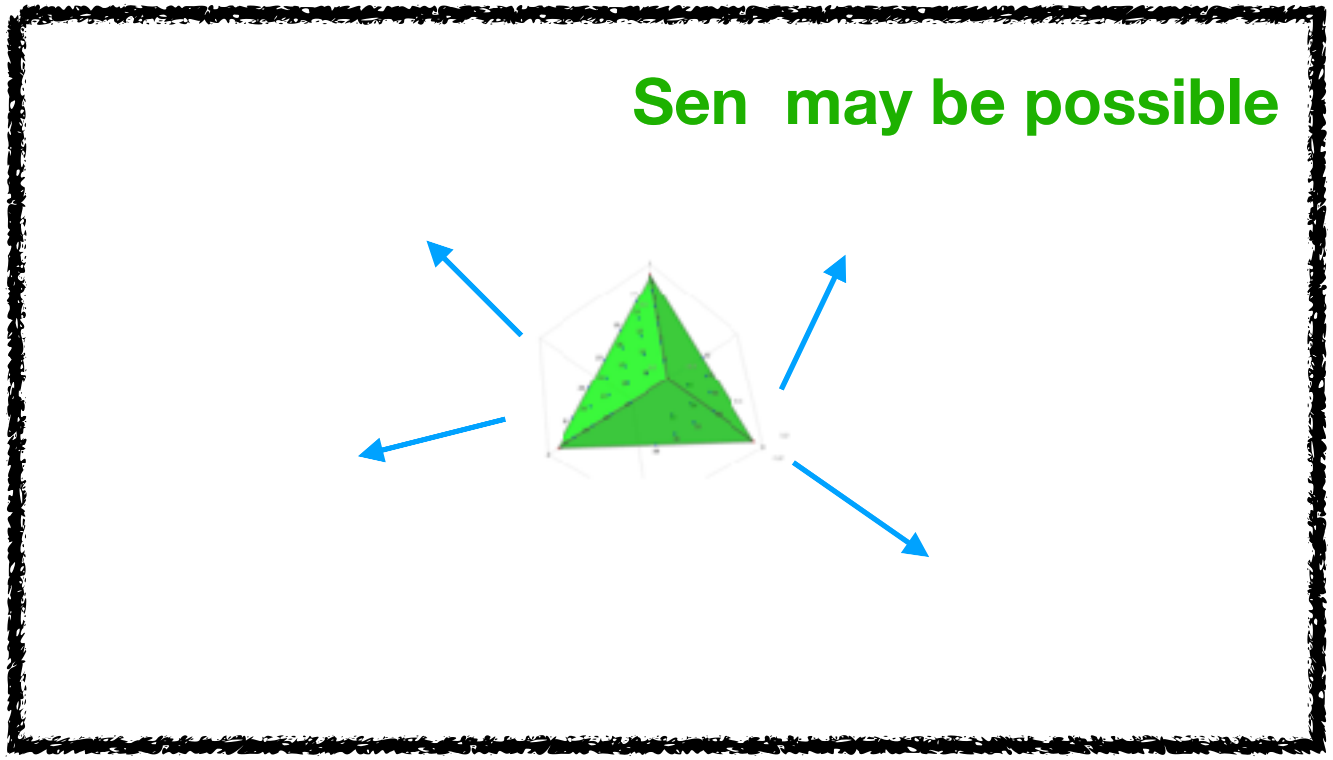
RL to Explore the Lamppost

why might this work? Sen-possible geometries **connected subset** of our ensemble, can start at FRST of 3d refl. poly.

No Sen limit!

The Game:

move: place tree
goal: stay in bounds
reward: 100 if in bounds
game over: out of bounds



still much to do, but want to show promising **initial results.**

(presented in time series of results, to emphasize fun)

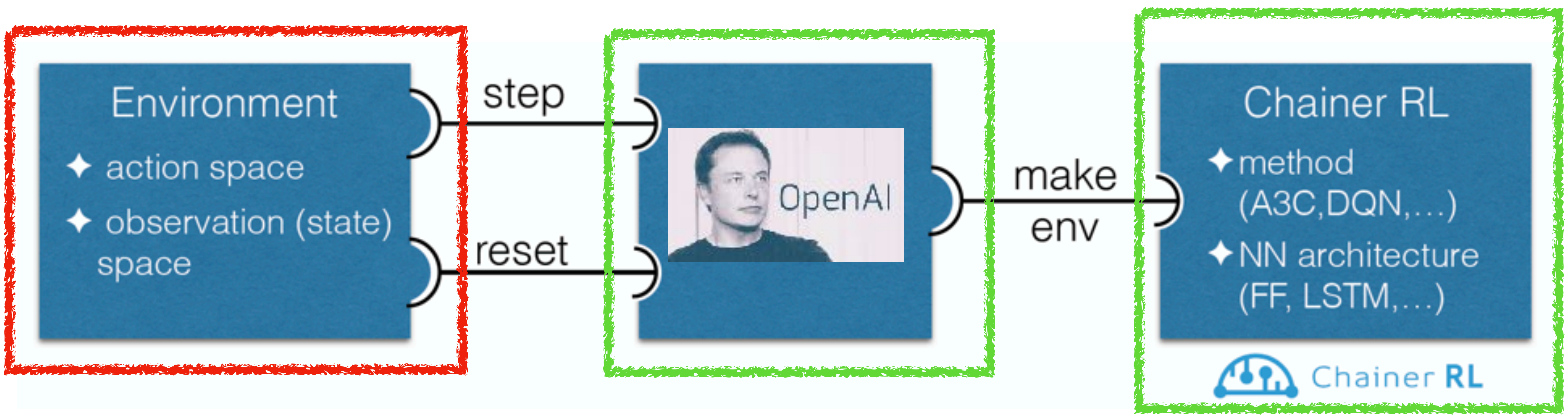
Implementation

model-free RL: want algorithms to work well regardless of environ.

means we can use CS-implemented algs!

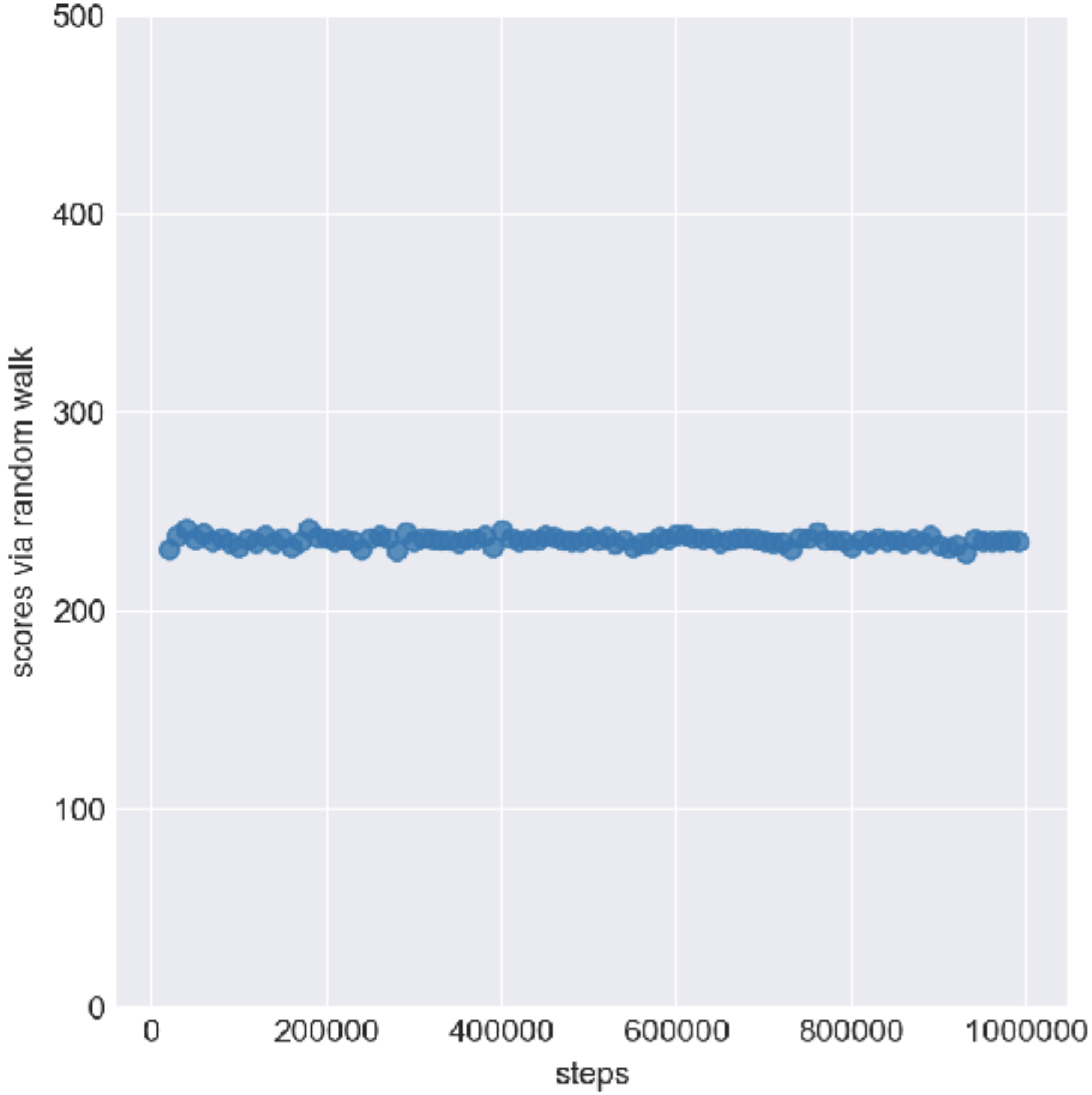
three modules:

- Open AI (Musk) defines what an environ is and how to interface.
- ChainerRL provides RL algorithms and NN architecture.
- **Physicists provide:** the environment. maintainsen-v0.



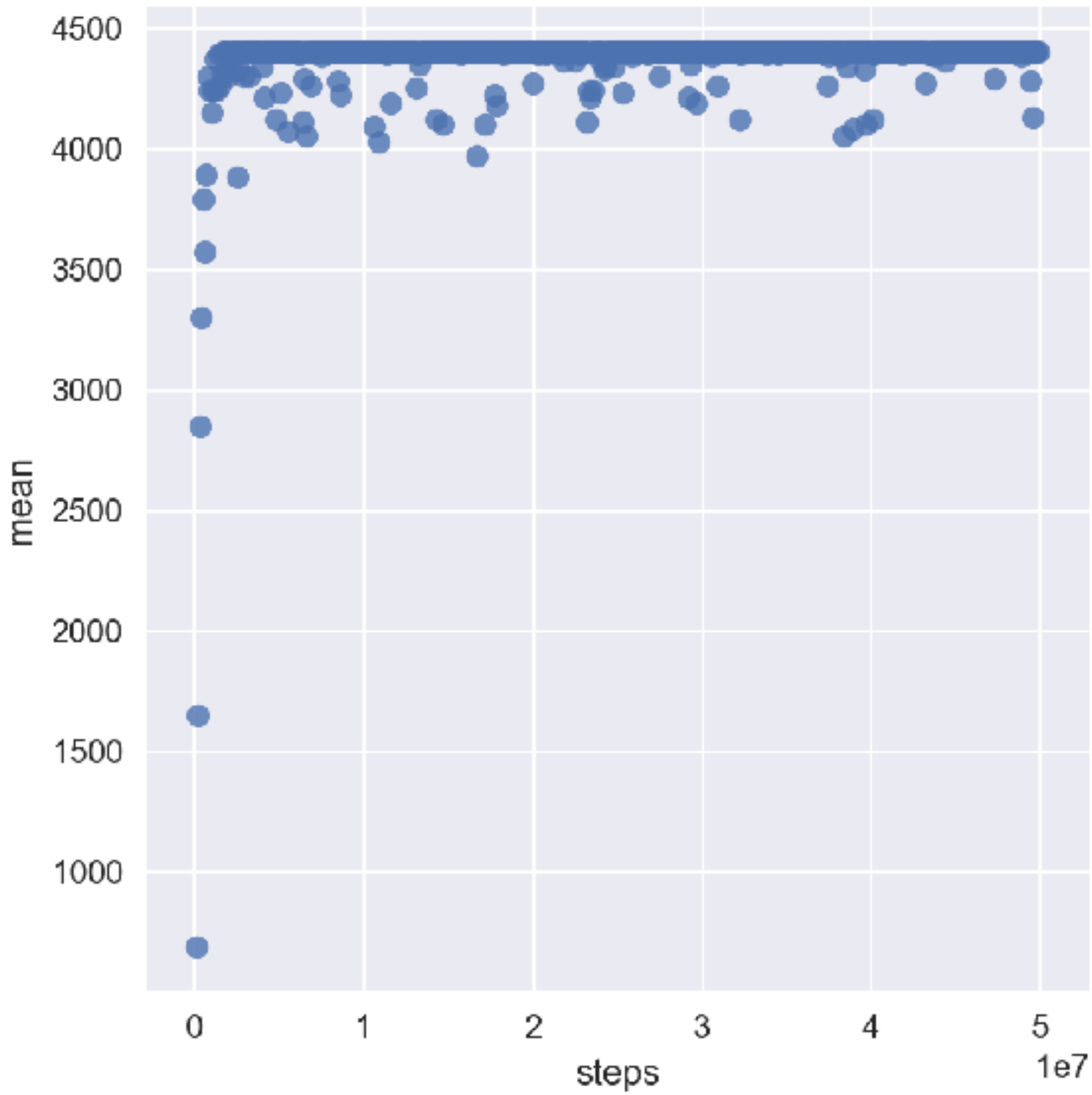
algorithm: asynchronous advantage actor-critic (A3C) [Minh et al 2016]
(parallel CPU, not GPU)

For Comparison: Random Walk



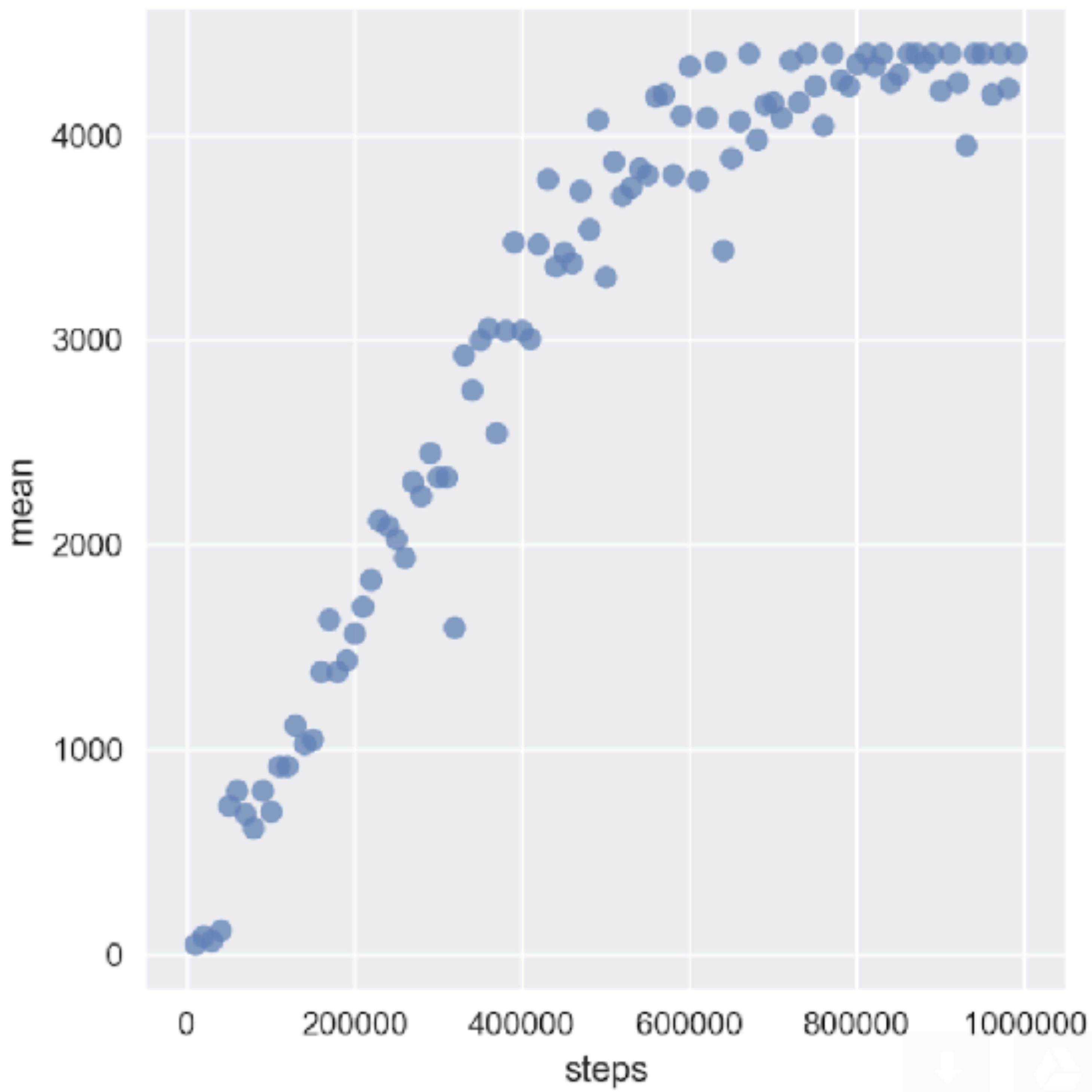
note scale: random walk takes 2-3 steps before NoSen

First Try: It Learns Quickly!



zoom in: decrease training time, increase eval interval

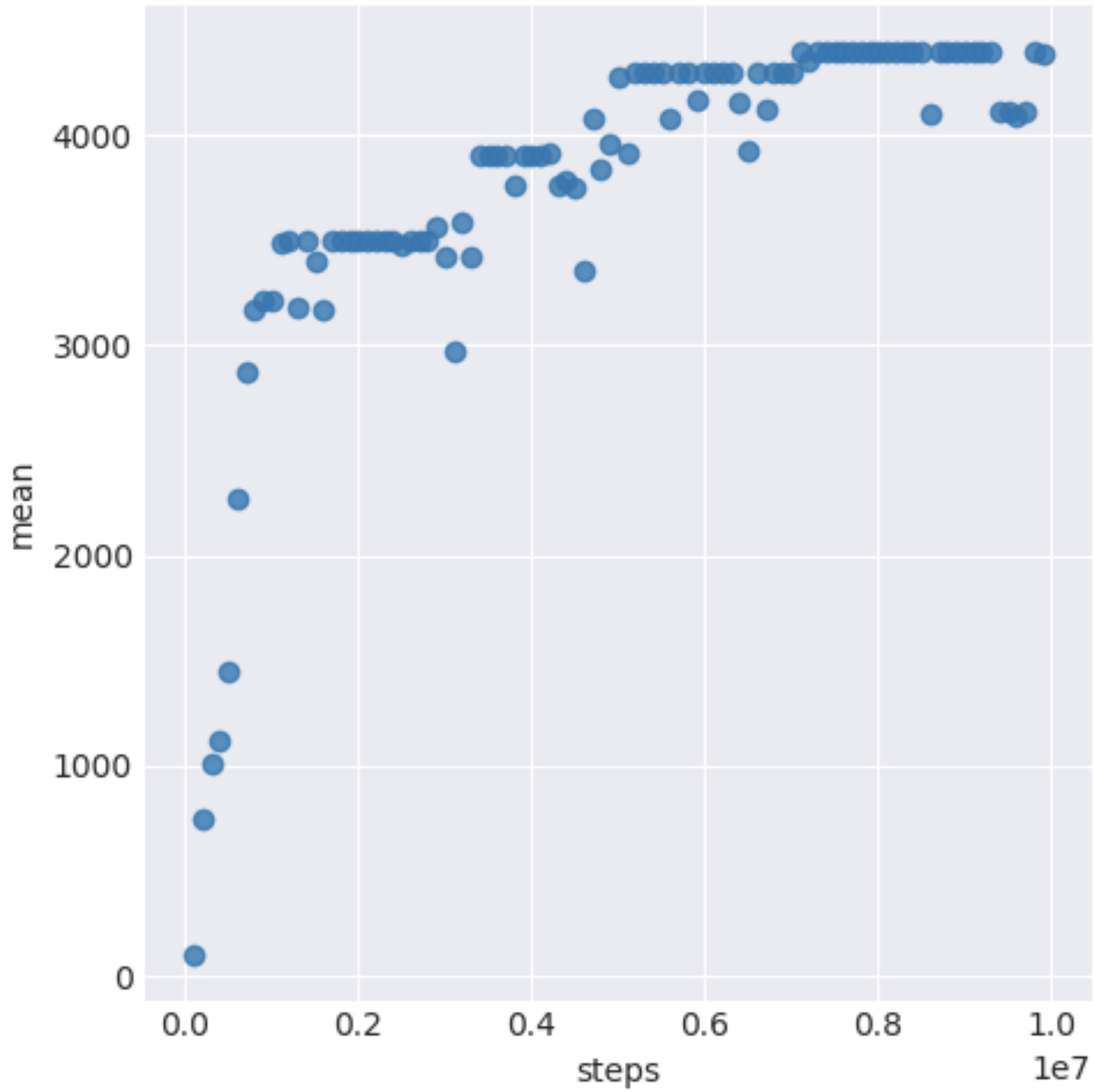
Second Try: See More Asymptote



much better, but can we tweak so it does better?

Third Try: Different Policy NN

use long short-term memory (LSTM) neural net

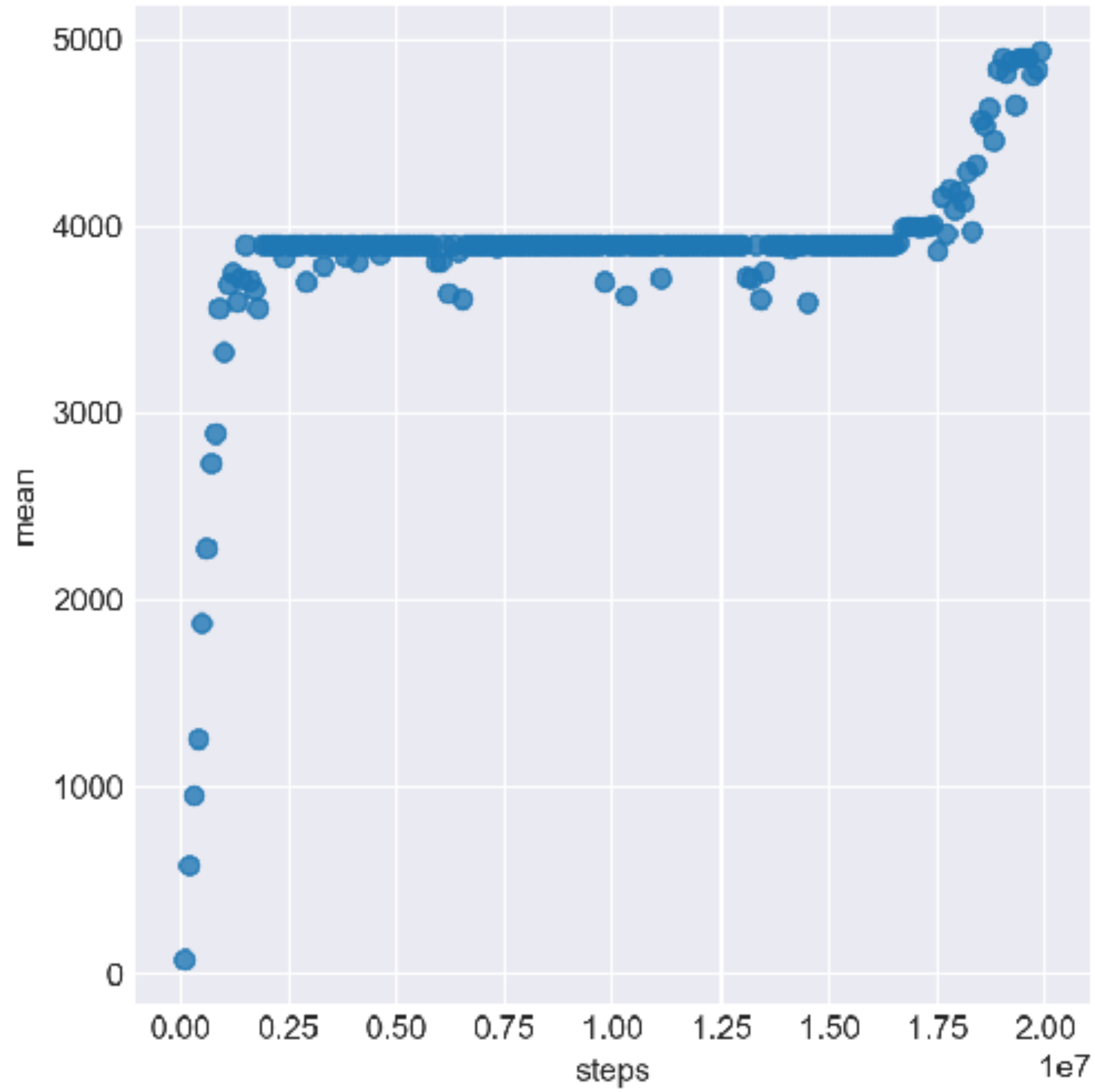


new feature: **four sharp plateaus**.

this is **punctuated equilibrium**, from evolution!

Fourth Try: Don't Give Up

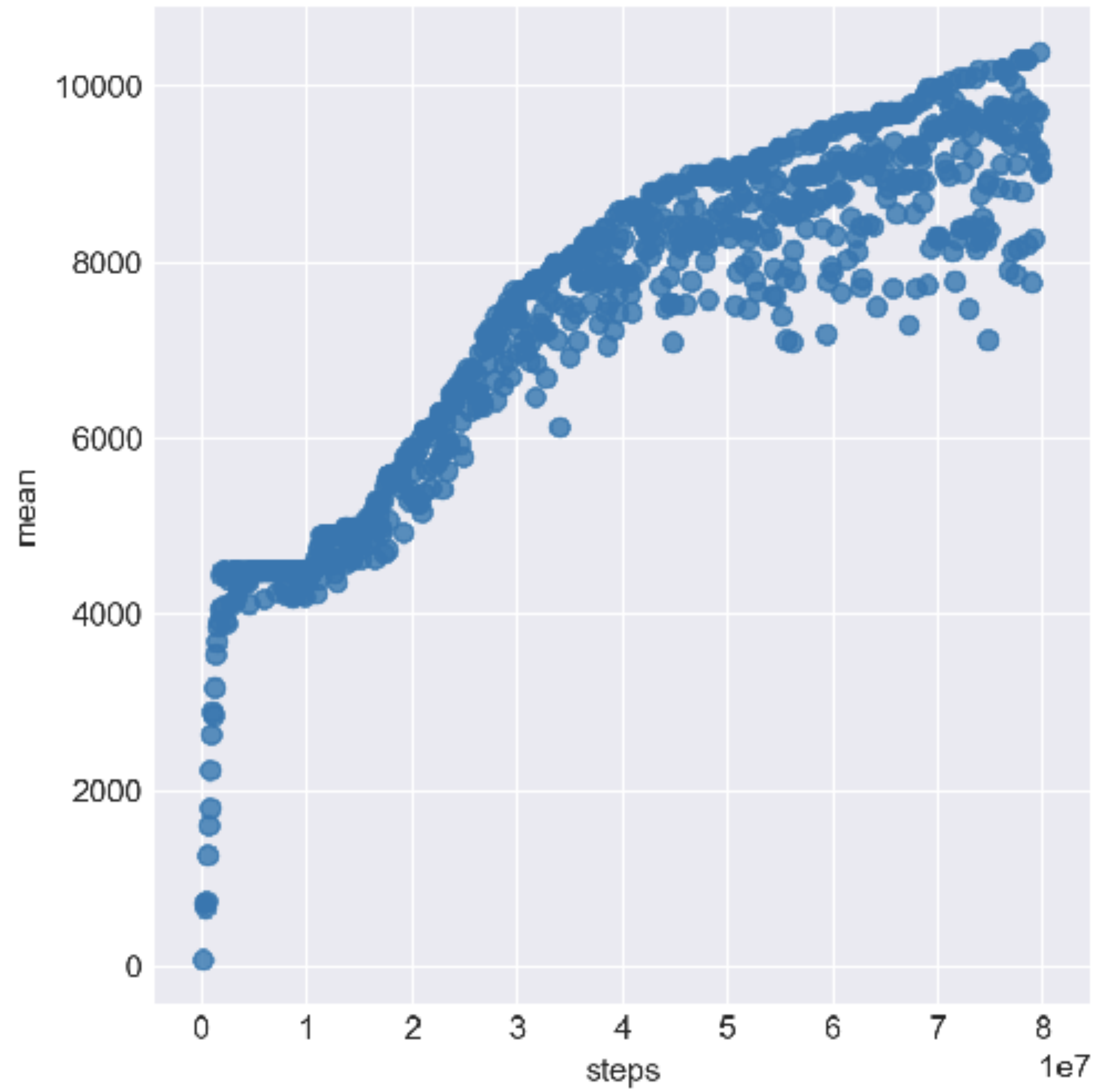
train a little longer, maybe it's got more juice in it.



work work work, keep on training.

Fifth Try: The Best Yet

and there's clearly still room to grow.



this is **training**, just phase 1. phase 2 and 3 to start soon.

Conclusions

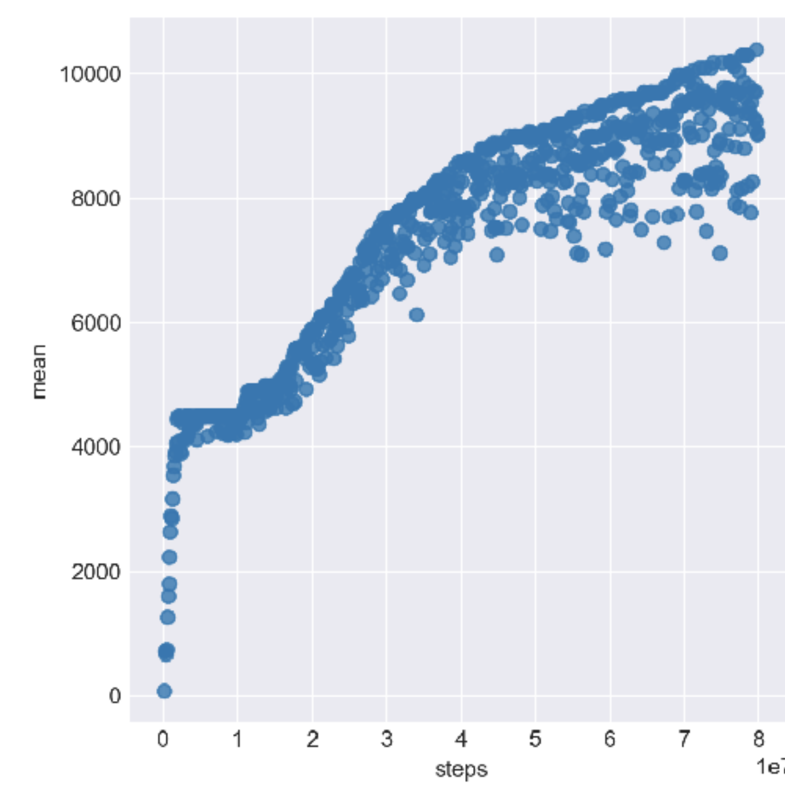
- The number of geometries in the landscape is **large**.
  We studied $4/3 \times 2.96 \times 10^{755}$ of them.

- NHC, large geometric gauge groups, and strong coupling are **universal** in this ensemble. **Matches Taylor-Wang**.

- **Anomaly detection:** rare phenomena are also interesting.

  Rigorous rare E6 result
  supervised ML + conjecture generation.

  Exploring IIB Lamppost
  reinforcement learning evolves AI agents.
  random walk: 2-3 steps.
  best AI so far: over 100 steps.



Extra Slides

AlphaGo Zero

"Mastering the game of Go without human knowledge."

Silver et al. (Google DeepMind), Nature Oct. 2017.

A long-standing goal of artificial intelligence is an algorithm that learns, tabula rasa, superhuman proficiency in challenging domains. Recently, AlphaGo became the first program to defeat a world champion in the game of Go. The tree search in AlphaGo evaluated positions and selected moves using deep neural networks. These neural networks were trained by supervised learning from human expert moves, and by reinforcement learning from self-play. **Here we introduce an algorithm based solely on reinforcement learning, without human data, guidance or domain knowledge beyond game rules.** AlphaGo becomes its own teacher: a neural network is trained to predict AlphaGo's own move selections and also the winner of AlphaGo's games. This neural network improves the strength of the tree search, resulting in higher quality move selection and stronger self-play in the next iteration. **Starting tabula rasa, our new program AlphaGo Zero achieved superhuman performance, winning 100–0 against the previously published, champion-defeating AlphaGo.**

Point: Go has 10^{172} states, therefore big, and for the task of playing excellently, superhuman progress achieved tabula rasa.

Is the String Landscape big?

- **Previous big landscape:**
  IIB flux vacua. Fix geometry, turn on fluxes.
  Flux estimates: $O(10^{500})$ Ashok, Denef, Douglas . . . $O(10^{272,000})$ Taylor, Wang

- **Emerging (?) big landscape:**
  Of topologically distinct geometries.
  Geometries: $4/3 \times 2.96 \times 10^{755}$ JH, Long, Sung $O(10^{3000})$ Taylor, Wang

- Logistical memory realities: Memory required for string geometries $\geq 10^{745}$ GB

  Memory required for flux vacua $\geq O(10^{272,000})$ GB.

- Logistical processing realities: (streaming algorithms?)

  Time required for streaming string geometries $\geq 10^{684}\ T_{\text{univ}}$

  Time required for streaming flux vacua $\geq O(10^{272,000})\ T_{\text{univ}}$,

How to handle a big landscape?

- **Algorithmic universality:** *universality derived not from a constructed set, but instead detailed knowledge of a concrete construction algorithm.*

- **Techniques from data science / AI for strings:**

  supervised machine learning. [He] [Krefl, Song] [Ruehle] [Carifio,JH, Krioukov,Nelson]

  (simple algs, neural nets, "predict")

  reinforcement learning / genetic algorithms: RL: [JH, Ruehle, Nelson] to appear. Genetic: [Abel, Rizos], [Ruehle]

  (DNN + psych, DNN + evolution, agents that learn, move, and "search")

  network science: ("connect") [Carifio, Cunningham, JH, Krioukov, Long, Nelson] [Taylor, Wang]

  topological data analysis: ("shape" of data) [Cole, Shiu] (for non-gaussianity)

  conjecture generation / intelligible AI: [Carifio,JH, Krioukov,Nelson]

  (use ML to generate conjectures, prove theorems. "rigorify".)

- **Vacuum selection:** maybe once we fully understand string theory, cosmological dynamics will allow us to ignore vast swaths of the landscape. (too hopeful?).

Full Set of E6 Anomaly Detection Slides

An E_6 Puzzle

- **Gauge group result:** dominated by $G_i \in \{E_8, F_4, G_2, A_1\}$

  (interesting: groups with only self-conjugate reps!)

- **Something SM-useful?** E6? SU(3)?

  - Simple conditions / probabilities for then not known. **JH, Long, Sung**
  - In random samples, prob ~ 1/1000.
  - When E6 arises in RS, on a distinguished vertex: (1,-1,-1).

- **Machine Learning:** **Carifio, Halverson, Krioukov, Nelson**

  Q: Can we train an ML model to accurately predict yes or no for E6 on (1,-1,-1)?

  Q: If so, can we learn how it makes its decision?

  in our paper: called **conjecture generation**.
  as a CS buzzword: **intelligible AI**.

Point: by using machine learning to generate conjectures, we may be able to take its numerical / empirical results and turn it into rigorous results.

Training the Model

- **Supervised machine learning:** given a large number of (input,output) pairs, **learn to predict** output given input, and then **test on unseen data**, see how well the model does.

- **Training data:**

  Input: (max height above v, # of such rays) for all v in polytope.
  Output: E6 on (1,-1,-1) or not.

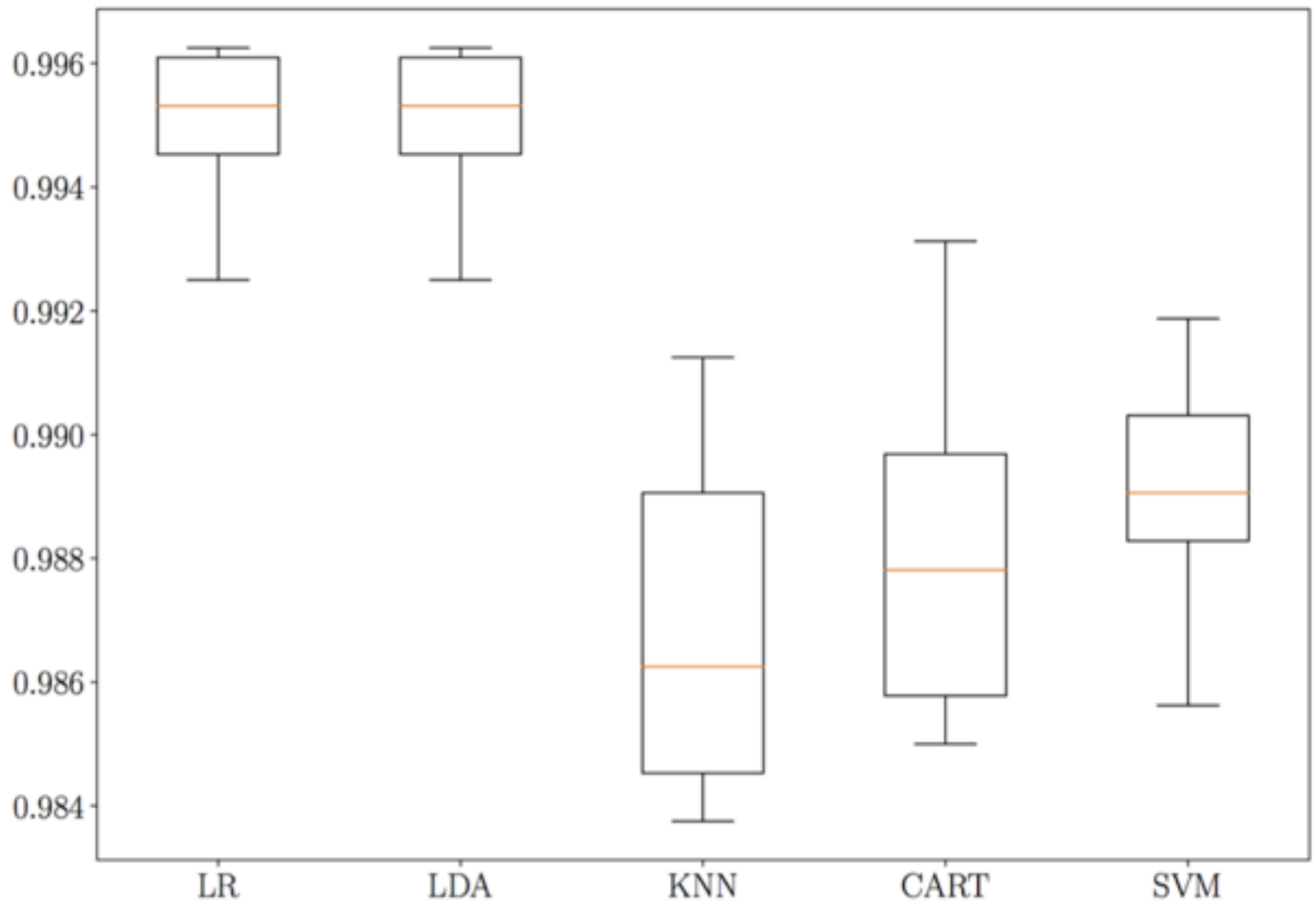
$$S_{a,v_1} := \{v \in V | v = av_1 + bv_2 + cv_3, \;\; a,b,c \geq 0\}$$

$$(a_{max}, |S_{a_{max},v}|) \quad \forall v \in \Delta_1^\circ \qquad \xrightarrow{A} \qquad E_6 \text{ on } v_{E_6} \text{ or not}$$

- **sklearn:** a very nice free Python package.

- **Training sample:** 10000 random with no E6, 10000 random with E6.

Evaluating the Model on Unseen Data



	LR	LDA	KNN	CART	SVM
50/50 Validation Set	.994	.994	.982	.987	.989
Unenriched Set	.988	.988	.981	.988	.983.

- **Displayed:** whisker plots of % accuracy with 10-fold cross validation.
- **Gold bar:** mean % accuracy.
- **Factor analysis:** only two of the variables really matter:

$$(a_{max}, |S_{a_{max}, v_{E6}}|)$$

Conjecture Generation

- **Organizing principle?** See what it gets right and wrong! (using the model trained with logistic regression.)

- **Observation:**
  amax = 5: always no
  amax = 4: usually no.

- **Initial Conjecture:**

Conjecture: If $a_{max} = 5$ for v_{E6}, then v_{E6} does not carry E_6. If $a_{max} = 4$ for v_{E6} it may or may not carry E_6, though it is more likely that it does.

a_{max}	$\lvert S_{a_{max},v_{E6}}\rvert$	Pred. for E_6 on v_{E_6}	Hyperplane Distance
4	5	No	0.88
4	6	No	0.29
4	7	Yes	-0.31
4	8	Yes	-0.90
4	9	Yes	-1.50
4	10	Yes	-2.09
4	11	Yes	-2.69
4	12	Yes	-3.28
4	13	Yes	-3.88
4	14	Yes	-4.47
4	15	Yes	-5.07
4	16	Yes	-5.67
4	17	Yes	-6.26
4	18	Yes	-6.85
4	19	Yes	-7.45
4	20	Yes	-8.04
4	21	Yes	-8.64
4	22	Yes	-9.23
4	23	Yes	-9.83
4	24	Yes	-10.42
5	1	No	7.34
5	2	No	6.75
5	3	No	6.15
5	4	No	5.56
5	5	No	4.96
5	6	No	4.37
5	7	No	3.78
5	8	No	3.18
5	9	No	2.59
5	10	No	1.99
5	11	No	1.40
5	12	No	0.80

Conjecture Refinement and Theorem

- **Use info from ML, think a bit, write down conjecture.**

Theorem: Suppose that with high probability the group G on v_{E_6} is $G \in \{E_6, E_7, E_8\}$ and that E_6 may only arise with $\tilde{m} = (-2, 0, 0)$. Given these assumptions, there are three cases that determine whether or not G is E_6.

a) If $a_{max} \geq 5$, $\tilde{m}$ cannot exist in Δ_g and the group on v_{E_6} is above E_6.

b) Consider $a_{max} = 4$. Let $v_i = a_i v_{E_6} + b_i v_2 + c_i v_3$ be a leaf built above v_{E_6}, and $B = \tilde{m} \cdot v_2$ and $C = \tilde{m} \cdot v_3$. Then G is E_6 if and only if $(B, b_i) > 0$ or $(C, c_i) > 0$ $\forall i$. Depending on the case, G may or may not be E_6.

c) If $a_{max} \leq 3$, $\tilde{m} \in \Delta_g$ and the group is E_6.

- **Key point:** ML-inspired focus on one particular variable, led quickly (< 24 hours) to a theorem once identified.

  "Back and forth" process, could be of broad applicability.

Probability and Checks

- **Probability computation:**

$$P(E_6 \text{ on } v_{E_6} \text{ in } T) = \left(1 - \frac{36}{82}\right)^9 \left(1 - \frac{18}{82}\right)^9 \simeq .00059128$$

computed using # appropriate edge trees relative theorem.

Result:

$$\text{Number of } E_6 \text{ Models on } T = .00059128 \times \frac{1}{3} \times 2.96 \times 10^{755} = 5.83 \times 10^{751}.$$

- **Check:** with 5 batches, 2 million random samples each.

$$\begin{aligned} \text{From Theorem} &: \quad .00059128 \times 2 \times 10^6 = 1182.56 \\ \text{From Random Samples} &: \quad 1183,\ 1181,\ 1194,\ 1125,\ 1195 \end{aligned}$$